

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

of

**Qian Zhang
Ya-Qin Zhang**

and

Wenwu Zhu

for

**Resource Allocation in Multi-Stream IP Network for
Optimized Quality of Service**

ATTORNEY'S DOCKET NO. MS1-589US

Resource Allocation in Multi-Stream IP Network for Optimized Quality of Service

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This non-provisional application claims priority to US Provisional Patent Application Serial No. 60/208,659, filed on May 31, 2000, which is incorporated herein by reference.

[0002]

TECHNICAL FIELD

The present invention relates to systems and methods for streaming media (e.g., video and audio) over a network, such as the Internet.

BACKGROUND OF THE INVENTION

[0003] With the success of the Internet and growing network resources, audio and video streaming is of enormous interest among Internet multimedia applications. Recently, tools such as RealPlay™ software and Microsoft NetShow® software have been developed to stream media content over a network. There remain, however, challenges to the streaming of media over the IP-based Internet due to issues such as the lack of a quality of service (QoS) guarantee, bandwidth variations, packet-losses, delays varying from time to time, and unknown network characteristics.

[0004] Since the Internet is a shared environment and does not micro-manage utilization of its resources, end systems are expected to be cooperative by reacting to congestion properly and promptly. As a result, overall utilization of the network remains high while

each flow obtains a fair share of resources. Unfortunately, many of the current commercial streaming applications do not behave in a network-friendly fashion.

[0005] The available bandwidth in the Internet fluctuates frequently in nature. Most conventional streaming applications are unable to perform quality adaptation as available bandwidth changes, especially quality adaptation among multiple streams. Thus, these conventional streaming applications do not make effective use of the bandwidth.

[0006] To date several schemes have been developed for QoS management such as resource reservation, priority mechanism, and application level control. Prior art QoS management schemes and other background information, referred to elsewhere in the document, are presented in the following publications, each of which is incorporated herein by reference in entirety:

[0007] 1. R. Braden, L. Zhang, S. Berson et al, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", *RFC 2205*, September 1997 ("hereinafter, "Braden et al.");

[0008] 2. R. Rejaie, M. Handley, and D. Estrin, "Quality adaptation for congestion controlled video playback over the Internet", *Proceedings of SIGCOM 99* ("hereinafter, "Rejaie et al. [SIGCOM]");

[0009] 3. R. Rejaie, M. Handley, and D. Estrin, "An end-to-end rate-based congestion control mechanism for realtime streams in the Internet", *Proceedings of INFOCOMM 99*, 1999 ("hereinafter, "Rejaie et al. [INFOCOMM]");

[0010] 4. T. Chiang and Y. Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling", *IEEE Trans. Circuits Syst. Video Technol.*, Feb. 1997 ("hereinafter, "Chiang et al.");

[0011] 5. D. Sisalem and H. Schulzrinne, "The loss-delay based adjusted algorithm: A TCP-friendly adaptation scheme", *Proceedings of NOSSDAV'98*, 1998 ("hereinafter, "Sisalem et al.");

[0012] 6. J. Padhye, V. Firoiu, D. Towsley and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation", *Proceedings of SIGCOMM'98*, 1998 ("hereinafter, "Padhye et al.");

[0013] 7. O. Verscheure, P. Frossard and M. Hamdi, "MPEG-2 video services over packet networks: joint effect of encoding rate and data loss on user-oriented QoS", *Proceedings of NOSSDAV 98*, 1998 ("hereinafter, "Verscheure et al.");

[0014] 8. A. Vetro, H. F. Sun and Y. Wang. "MPEG-4 rate control for multiple video objects". *IEEE Trans. Circuits Syst. Video Technol.*, Feb. 1999 ("hereinafter, "Vetro et al."); and

[0015] 9. M. Eckert and J. I. Ronda. "Bit-rate allocation in multi-object video coding". ISO/IEC JTC1/SC29/WG11 MPEG98/m3757, Dublin, Ireland ("hereinafter, "Eckert et al.).

[0016] Among the foregoing QoS management schemes, resource reservation for supporting a certain QoS level, which was proposed by Braden et al., is the most straightforward approach. However, since it is difficult to know the characteristics of a stream in advance, one may tend to over-allocate resources in order to guarantee the requested QoS level, leading to network under-utilization. Besides that, the most challenging issue for the resource reservation mechanism is that it is both difficult and complex to implement and to deploy.

[0017] In priority mechanisms, different data packets or streams are labeled with different priorities and thereby treated differently at the network routers. While this approach is simple, the exact mechanism for setting the priority levels, the router mechanism for controlling these levels, and the actual gain are unclear.

[0018] In application level control scheme, the QoS is controlled by adapting the sender transmission rate as was taught by Rejaie et al. [*SIGCOM*] and Chiang et al. Most of the control algorithms, however, randomly distribute resources among multiple streams without a global coordination mechanism. In order to employ a global coordination scheme, traffic control is usually adopted. There are several TCP-friendly rate adjustment protocols that have been reported recently. It has been proposed that transport protocols, including those taught by Rejaie et al. [*INFOCOMM*] and Sisalem et al. who teach transport protocols where the throughput of a long-lived TCP connect is calculated based on the TCP characterization. However, Padhye et al. demonstrated that the above approaches to calculating the throughput are not accurate in cases where the packet-loss rate is higher than five percent (5%). Since this approach does not account for retransmission timeouts, it usually overestimates the throughput of a connection as the packet-loss rate increases.

[0019] Other challenges to streaming video include network bandwidth adaptation, media adaptation, and error resilience. Network bandwidth adaptation deals with dynamic network bandwidth estimation, while media adaptation controls the media bit rate according to the network conditions. Error resilience refers to the ability to track, localize and recover transmission errors.

[0020] Without bandwidth adaptation, video transmission tends to compete unfairly with other TCP traffic, causing network congestion and resulting in a lower transmission rate for other TCP traffic. In addition, congestion collapse occurs when the aggregate bandwidth of the media traffic exceeds the network capacity. In order to dynamically adjust the transmission rate while co-existing with other TCP-based applications, several congestion control protocols have been proposed to adapt the sending rate in such a way that congested bandwidth is shared “fairly” with TCP applications. The majority of these protocols are based on TCP characterizations. Specifically, in the absence of retransmission time-outs, the steady state throughput of a long-lived TCP connect is given by:

$$\text{Throughput} = \frac{C}{R * \sqrt{p}}, \quad (1)$$

where C is a constant that is usually set to either 1.22 or 1.31, depending on whether the receiver uses delayed acknowledgments, R is the round trip time experienced by the connection, and p is the expected number of window reduction events per packet sent.

Since Equation (1) does not account for retransmission timeouts, it usually overestimates the connection throughput as packet-loss rate increases. It is has been reported that Equation (1) is not accurate for packet-loss rates higher than 5%.

[0021] MPEG-4 is an object-based coding standard in which a visual scene typically has several video objects (VOs), each characterized by its shape, motion, and texture. The VOs are coded into separate bit streams that can be individually accessed and manipulated. The composition information is sent in a separate stream. To date several prior art rate control algorithms have been proposed. One such rate control algorithm was proposed by Chiang et al. for a single VO using a quadratic rate-quantizer model as the

baseline rate control scheme used in the MPEG4 standard. This rate control scheme was extended by Vetro et al. to multiple video objects (MVOs). As taught by both Chiang et al. and Vetro et al., the total target bit rate for all objects is controlled by a “joint-buffer” and allocated proportionally to the motion, size, and square of MAD (mean absolute distortion). For MVOs, Eckert et al. taught several straightforward approaches based upon the video object importance level. Although these approaches allow different objects to be encoded at different frame rates, when put in one scene, these objects with different frame rates may cause a break in the composition information and can result in unacceptable video reconstruction.

[0022] When MPEG-4 video is transported over the Internet, all the above rate control schemes could not work well since they do not adapt to network bandwidth and packet-loss conditions that vary from time to time. Thus, the available resources could not be efficiently utilized and sometimes they may suffer from heavy congestion.

[0023] It would be an advance in the art to devise a multimedia streaming TCP-friendly transport protocol that can adaptively estimate the network bandwidth and smooth the sending rate. It would also be an advance in the art to devise a global resource allocation control mechanism that maximizes the quality of AV streams delivered across fairly congested connections, where bits are allocated dynamically according to the media encoding distortion and network degradation. With respect to multiple video objects, it would be an advance in the art to devise a rate control scheme that uses such a multimedia streaming TCP-friendly protocol while minimizing the overall distortion under the constraint that the total rate for all objects is upper-bounded by a target bit rate. Finally,

an advance in the audiovisual streaming art would be achieved by minimizing the end-to-end distortion for a given network traffic condition and picture quality requirement.

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
224

BRIEF SUMMARY OF THE INVENTION

[0024] The resource allocation architecture allocates resources among multiple media streams over a network (such as the Internet) to achieve optimized end-to-end quality of service. Examples of the media include audio, MPEG/H.26x compressed video, mesh, compressed image, and so forth. The resource allocation architecture copes with packet drops and fluctuating available bandwidth. From the network congestion control viewpoint, the architecture allows a multimedia streaming TCP-friendly protocol, which is suitable for multimedia streaming. From the visual quality control viewpoint, the architecture provides a way to maximize the visual quality for all streams based on the variable available bandwidth and the characteristics of each of the multiple media streams. The resource allocation architecture thereby enables the user to gain globally optimized quality at the client side.

[0025] A method is disclosed for transmitting a mixed media data stream, including audio and multiple video objects (MVOs), between a sender and a receiver through a connection over a network. The transmission characteristics of the connection between server and receiver are monitored and the available bandwidth is estimated at the sender based upon the transmission characteristics of the connection monitored at the receiver side. A global buffer is allocated for the mixed media data stream to be transmitted from the sender to the receiver as a function of the estimated available bandwidth at the sender. A portion of each Video Object Plane (VOP) in the global buffer is pre-encoded with respect to a quantization parameter (QP) of the VOP. The VOP in the global buffer is then encoded based on the QP. An update is performed to a rate distortion model based upon the QP and a frame skipping function is performed after the VOP encoding. The sender then

transmits the encoded video object plane in the global buffer at a regulated sender transmission rate from the sender as a function of the estimated available bandwidth at the sender.

[0026] In one implementation, a multimedia streaming TCP-friendly transport protocol is used to adaptively estimate the network bandwidth and smooth the rate at which a server sends streaming data to a receiver. A global resource allocation control mechanism is also implemented. This allocation control maximizes the quality of audio-visual (AV) streams delivered across congested connections, where bits are allocated dynamically to a global buffer for transmission over a network. This dynamic bit allocation is made in a manner that minimizes the media encoding distortion and network degradation.

[0027] In an implementation dealing with multiple video objects, a rate control scheme uses a TCP-friendly protocol that minimizes the overall distortion under the constraint that the total rate for all objects is upper-bounded by a target bit rate. The TCP-friendly protocol is used to obtain network characteristics, including packet-loss rate, delay and jitter. Network characteristics are used to estimate available network bandwidth and to make adjustments to the sending rate in a smooth, non-erratic manner. The sending rate control scheme implemented herein allows the target bit rate and global buffer size to be dynamically adapted to the estimated available network bandwidth. Additionally, the packet-loss rate is taken into account when calculating the overall distortion.

[0028] In each implementation, end-to-end distortion is minimized for network traffic conditions and picture quality requirements.

[0029]

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a general architecture where several continuous media servers play back AV streams for heterogeneous clients on demand over the Internet, where each server is able to support a large number of requests simultaneously, and where each client is able to request services from different servers.

[0030] Figure 2 depicts an end-to-end framework for media streaming over the Internet.

[0031] Figure 3 depicts the Gilbert Model which is able to capture the dependence between consecutive packet-losses, and wherein “1” is the received state and “0” is the loss state.

[0032] Figure 4 illustrates a topology used in a simulation of an implementation of a resource allocation scheme.

[0033] Figures 5 and 6 show simulation results of the throughput and sending rate, respectively, for six (6) different connections with their respective times.

[0034] Figure 7 is a plot of peak signal to noise ratio (PSNR) per frame of different video objects, wherein the plot is illustrated with and without the disclosed resource allocation control scheme.

[0035] Figure 8 is a plot of the PSNR per frame using two of the four objects in the MPEG-4 test sequence “Coast Guard”, wherein the plot is illustrated with and without the disclosed resource allocation control scheme.

[0036] Figure 9 shows a block diagram of the network resource allocation scheme disclosed herein for multiple video objects (MVOs).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0039] The following discussion describes a system and method to allocate network resources among multiple media streams. These network resources are allocated within a client-network-server framework (such as the Internet) in a TCP-friendly manner to achieve optimized end-to-end quality of service. Simulation results are given that depict an implementation of the allocation of network resources to yield optimized end-to-end quality of service. A final discussion is then presented of an implementation of a network adaptive rate control scheme, including the results of a simulation of the same.

[0040] 1. FRAMEWORK FOR INTERNET AUDIO-VISUAL STREAMING

In media streaming, multiple servers and multiple clients are usually employed in the same session. Figure 1 depicts a general architecture where several continuous media servers play back AV streams for heterogeneous clients on demand, in which $Server_1$, $Server_i$, $Server_j$, $Server_n$ are in communication with $Client_1$, $Client_p$, $Client_q$, and $Client_m$ through the Internet. Each server is able to support a large number of requests simultaneously and each client is able to request services from different servers.

[0041] Figure 2 depicts an implementation of an end-to-end framework for media streaming over the Internet seen at reference numeral 220. End-to-end transport control is adopted by using a Multimedia Streaming TCP-Friendly transport Protocol (MSFTP) seen at reference numerals 212 and 222. The MSFTP protocol is a rate-based TCP-friendly protocol that continuously monitors the connection between sender and receiver, which are depicted in Figure 2, respectively, as Server 210 and Client 230. The results of the continuous monitoring of the connection are then used to regulate the bit transmission rate of Server 210.

[0042] A Quality of Service (QoS) Adaptation and Global Resource Allocation Control module 214 adjusts the quality of the bit stream that is transmitted from server 210 to client 230 through Internet 220. Module 214 periodically estimates the available bandwidth from the MSTFP protocol as derived from module 212. Combining this information with the media characteristics of media streams Video_m, Video_n, and Audio_k, module 214 adjusts the quality of the total transmitted streams by resource re-allocation in a Global Buffer Control module 216 at Server 210. Data occupying the global buffer is transmitted from Server 210 in the MSFTP protocol at module 212 through Internet 220 to Client 230.

[0043] Client 230 processes the MSFTP protocol at module 232 and passes the media to a demultiplexer module seen in Figure 2 at reference numeral 234. Demultiplexer module 234 demultiplexes the combined stream into original media types for decoding at Video_m decoder, Video_n decoder, and Audio_k decoder. The output from the video and audio decoders are mixed at a Media Mixer module 236 and output to an output device 238, such as a personal computer at Client 230 having a display device and having a sound card with associated speakers.

[0044] A discussion of the MSFTP protocol follows. Then, two exemplary implementations that use the MSFTP protocol are set forth. The first implementation is a resource allocation scheme for audio-visual streaming that is discussed in conjunction with the example illustrated in Figure 2. The second implementation is a network adaptive sending rate control scheme that is discussed in conjunction with the example illustrated in Figure 9. Both implementations will be discussed conceptually and in a presentation of results obtained by respective experimentations of each.

**[0045] 2. MULTIMEDIA STREAMING TCP-FRIENDLY
TRANSPORT PROTOCOL (MSTFP)**

To transmit the video and audio streams over the Internet, the characteristics of packet-loss and bandwidth fluctuation should be taken into account. The multimedia streaming TCP-friendly protocol (MSTFP) intends to minimize the number of future packets likely to be dropped and then smoothes the sending rate. Integral to the MSTFP protocol is the underlying TCP model taught by Padhye et al. which allows the sender to change its sending rate based on the observed packet-loss rate, round trip times, and retransmission time-outs. The two main characteristics of the MSFTP protocol are “TCP-friendliness” and “rate smoothness”.

[0046] The MSTFP protocol involves sender-side data packets and receiver-side control packets. The sender sends data packets to the receiver at a certain rate. The header of the sender-side packet includes the packet sequence number, a timestamp indicating the time when the packet is sent (STI), and the size of the sending packet. The receiver sends feedback to the sender at regular intervals. The receiver-side packet includes the time interval that the packet spent in the receiver side (ΔRT), the timestamp of the packet sent from the sender (STI), the estimated packet-loss rate, and the rate at which data is received. Based on the receiver’s feedback, the sender uses the TCP model taught by Padhye et al. to adjust its sending rate in a TCP-friendly manner.

[0047] What follows is a discussion of packet-loss rate estimation, round trip time estimation (RTT), retransmission time-out estimation (TO), available bandwidth estimation, and sending rate adjustment, each of which apply to the MSTFP protocol. The MSTFP protocol is preferred for the resource allocation implementation for audio-visual streaming and for the network adaptive sending rate control implementation.

[0048] A. Packet-loss rate estimation

A two-state Markov Model (Gilbert model) is used to model the packet-loss in the network. Figure 3 depicts the Gilbert Model, wherein “1” is the received state and “0” is the loss state. This model is able to capture the dependence between consecutive losses. Network packets can be represented as a binary time series, $\{x_i\}_{i=1}^n$, where x_i takes 1 if the i th packet has arrived successfully and 0 if it is lost. The current state, X_i , of the stochastic process depends only on the previous value, X_{i-1} . The transition probabilities between the two states are calculated as follows:

$$p = P[X_i = 1 | X_{i-1} = 0] \text{ and } q = P[X_i = 0 | X_{i-1} = 1]. \quad (2)$$

[0049] The maximum likelihood estimators of p and q for a sample trace are:

$$\hat{p} = n_{01}/n_0 \text{ and } \hat{q} = n_{10}/n_1, \quad (3)$$

where n_{01} is the number of times in the observed time series when 1 follows 0 and n_{10} is the number of times when 0 follows 1. n_0 is the number of 0s and n_1 is the number of 1s in the trace.

[0050] The probability of the event that the trace is in the loss state is given by:

$$P_L = \frac{\hat{q}}{\hat{p} + \hat{q}}. \quad (4)$$

[0051] To obtain an accurate and stable measurement of the packet loss rate, the scheme uses a filter that weights the n ($n = 8$ in this implementation) most recent packet loss rate values in such a way that the measured packet loss rate changes smoothly. Thus, the estimated packet loss rate in the i th time interval is calculated as:

$$P_{L,i} = \sum_{j=0}^{n-1} (w_j \times \overline{P_{L,i-j}}) \quad (5)$$

where $\overline{P_{L,i-j}}$ is the measured packet loss rate in the (i-j)th time interval.

[0052] It can be seen that the weighting parameters values, w_j , used in calculating the packet loss rate determines the protocol's speed in responding to changes in the level of congestion. Two sets of weighting parameters are defined in this implementation, which are denoted as WS1 and WS2, respectively.

	W0	W1	W2	W3	W4	W5	W6	W7
WS1	1.0	1.0	1.0	1.0	0.8	0.6	0.4	0.2

	W0	W1	W2	W3	W4	W5	W6	W7
WS2	1.2	1.2	1.0	1.0	0.8	0.5	0.3	0.1

[0053] In the above two parameter sets, WS1 is used in the default cases. When the most recent packet loss rate is less than the half of the calculated packet loss rate, WS2 is adopted to concentrate more of the relative weight on the most recent packet loss rate.

[0054] B. RTT and TO estimation

Based on the receiver's feedback packet, the sender can adjust the current RTT using the following equation:

$$RTT = \alpha \times \overline{RTT} + (1 - \alpha) \times (now - ST1 - \Delta RT), \quad (6)$$

where \overline{RTT} is the current round trip time, RTT is the estimated round trip time, now is the timestamp indicating the time at which the packet was received in the sender, and α is a weighting parameter that is set to 0.75 to smooth the variation for the purposes of this implementation.

[0055] After calculating the RTT, retransmission timeout can be defined as:

$$TO = RTT + (k \times RTTVAR), \quad (7)$$

where k is a constant that is recommended to take the value 4, RTT is the estimated round trip time of the sender packet from the sender to the receiver, and RTTVAR is a smoothed estimate of the variation of RTT. RTTVAR in Eq. (7) can be represented as:

$$RTTVAR = \alpha_2 \times \overline{RTTVAR} + (1 - \alpha_2) \times |RTT - (now - ST1 - \Delta RT)|. \quad (8)$$

In Eq. (8), \overline{RTTVAR} is the current variation in the round trip time of the sender packet from the sender to the receiver (RTT), α_2 is a weighting parameter that is preferably set to 0.75, and RTTVAR is a smoothed estimate of \overline{RTTVAR} ..

[0056] C. Available bandwidth estimation

After the above procedures, the sender can estimate the present available network bandwidth as taught by Padhye et al. who calculate the present available network bandwidth as follows:

$$rcvrate = \frac{PacketSize}{RTT \times \sqrt{2P_L/3} + 3 \times TO \times P_L \times \sqrt{3P_L/8} \times (1 + 32P_L^2)}, \quad (9)$$

where *PacketSize* is the size of the sending packet and *TO* is a Time Out period that indicates the length of a time interval before which a sender should wait before re-transmitting a packet of data. As taught by Padhye et al., the present implementation proposes a TCP protocol that calculates *TO* without re-transmitting the data.

[0057] D. Sending rate adjustment

[0058] After estimating the available network bandwidth, the sender can dynamically adjust its sending rate. TCP congestion control is based on Additive Increase/

Multiplicative Decrease (AIMD), which is independent of the lost fraction and adjusting interval. An alternate way to change the sending rate related to current packet-loss is described as:

(10)

```

if (rcvrate >  $\overline{currate}$ )
    multi = (now - lastchange)/RTT
    constraint multi from 1 to 2
     $currate = \overline{currate} + (\text{PacketSize}/\text{RTT}) \times \text{multi}$ 
else
     $currate = \beta \times \text{rcvrate} + (1 - \beta) \times \overline{currate}$  ,

```

where $\overline{currate}$ is the present sending rate, $currate$ is the updated sending rate, $lastchange$ is the timestamp indicating the time at which last adjustment occurred, and β is the weighting parameter that is set to 0.75 in this implementation. An advantage of the above scheme is that the sending rate can be increased dramatically to the available bandwidth, and in the meanwhile the rate can be decreased smoothly. In other words, the MSTFP protocol has less variation in the transmission rate and is less sensitive to random loss caused by channel error.

[0059] 3. RESOURCE ALLOCATION FOR AUDIO-VISUAL STREAMING

What follows is a discussion of an implementation of a resource allocation scheme for audio-visual streaming and a presentation of the results of simulations of the same.

[0060] A. Discussion

Dynamic resource allocation is preferable for distributed multimedia systems that support application-level control. In the example illustrated in Figure 2, dynamic resource allocation can be implemented by integrating global buffer control 216 with Quality of Service Adaptation and Global Resource Allocation Control 216 using dynamic network bandwidth estimation. Dynamic network bandwidth estimations are made using the changing characteristics of Internet 220 learned at modules 212 and 232 using the MSFTP protocol. One of the most challenging tasks of dynamic resource reallocation is global coordination of the feedback information from multiple streams.

[0061] Different applications, such as file-transfer, web browsing, and AV streaming, have different tolerances that can be mismatched between the sending rate and the network bandwidth. For example, file-transfer has no real-time constraints and is tolerable to delay. The difference between the sensitivities to human aural and visual systems indicates the audio and video should be handled differently when adverse conditions arise, thereby affecting the playback of media streams. It is well known that the aural sense is more sensitive to disturbances than the visual sense. Therefore, it is appropriate to assign higher priority to audio data than video data. If data can be discarded when congestion occurs in the network, it is preferable to discard video data first.

[0062] In general, the objective of resource allocation is to minimize the overall distortion under the total bit-rate constraint. If the sending rate of the i th media stream is denoted as r_i , the distortion of i th media as d_i , and the quality-impact parameter of this stream as α_i , then a minimization of the distortion can be expressed as:

$$\text{Minimize } D = \sum_i \alpha_i \times d_i, \text{ subject to } R = \sum_i r_i \leq R_T, \quad (11)$$

where R_T is the total bit budget for the current time instant obtained from the MSTFP protocol, i.e.,

$$R_T = \sum_i currate_i. \quad (12)$$

[0063] In the resource allocation scheme disclosed below, the main focus is on global buffer control and bandwidth allocation, where a global buffer is used to control the bit rate allocation. Using the MSTFP protocol, the current available network bandwidth R_T , can be estimated. Together with the number of bits spent in the global buffer in the previous time instant B_{prev} , the size $R_{old}/2$, and the occupancy of the global output buffer W_{prev} in the previous time instant, the target rate and global buffer can be updated for each time interval. The output target rate of the joint buffer control is allocated among each stream to yield the target bit for an individual object. As the buffer size $R_{old}/2$ is changed to $R_T/2$, the occupancy of the buffer is changed as follows:

$$W_{cur} = \max(((W_{prev} + B_{prev}) \times R_T / R_{old} - R_T / F), 0), \quad (13) \quad \text{where "F" is the video frame rate (e.g., 20 frame/second, 30 frame/second, etc.)}$$

[0064] To achieve the global minimal distortion among the multiple streams, it is desirable to define the appropriate function for rate and distortion.

[0065] To predict the number of bits that is used to encode the current stream for video objects, a quadratic function calculation is used that is taught by Chiang et al., and is given by :

$$r_i = \frac{p_{i1} \times MAD_i}{Q_i} + \frac{p_{i2} \times MAD_i}{Q_i^2} \quad (14)$$

[0066] Verscheure et al. teach that packet-loss has an impact on video quality. For video distortion measuring, the present implementation considers the packet-loss that occurs

during video transmission. The video quality will drop quickly as the packet-loss rate and the sending rate go up.

[0067] To take the packet-loss into account, the implementation uses another quadratic function to estimate the distortion that is given by:

$$d_i = q_{i1} \times Q_i + q_{i2} \times Q_i^2 + q_{i3} \times r_i \times P_{L_i}. \quad (15)$$

[0068] In the audio case, one could still use the above distortion and rate control algorithms, respectively, in Eqs. (14) and (15). In Eq. (14), MAD_i is the energy for the audio case.

[0069] Having the information of the rate and distortion, distortion can be minimized by Lagrange optimization to obtain the appropriate quantization parameters. These estimated quantization parameters are used in the encoding process.

[0070] B. Simulation Results

MPEG-4 streams with multiple video objects were used in a simulation to demonstrate the effectiveness of an implementation of the network resource allocation scheme for audio-visual streaming. MPEG-4 is an object-based video coding standard in which a visual scene is typically composed of Video Objects (VOs). Each VO is individually coded to give rise to an elementary bit stream that can be individually accessed and manipulated. The composition information is sent in a separate stream. For multiple video object streams, foreground objects receive most attention from the viewer, while background objects are of less importance. In addition, foreground objects usually change rapidly, while background objects change slowly. Therefore, they may have different impact on the overall video quality.

[0071] In the simulation, the Network Simulator (NS) version 2 software was used to study the MSTFP protocol and the resource allocation mechanism for the MPEG-4

multiple video objects. The main tasks of the simulation were to select appropriate network topologies, effectively model the background traffic, and define proper performance metrics.

[0072] Figure 4 is a topology used in the simulation. The topology has a single shared bottleneck link. The sources or senders are on one side of the link in Fig. 4 and the receivers are on the other side. All links except the bottleneck link are sufficiently provisioned to ensure that any drops/delays that occurred are only caused by congestion at the bottleneck link. All links are drop-tail links. In the simulation, the background traffic has infinite-duration TCP-like connections and infinite-duration real-time adaptive protocol (RAP) connections that were proposed by Rejaie et al. [*SIGCOM*] to deliver real-time traffic over Internet.

[0073] As used herein, “friendliness” metrics are used as follows. k_m denotes the total number of monitored MSTFP protocol connections and k_t denotes the total number of monitored TCP connections. The throughputs of the MSTFP protocol connections are further denoted as $T_1^m, T_2^m, \dots, T_{k_m}^m$ and the throughputs of the TCP connections as $T_1^t, T_2^t, \dots, T_{k_t}^t$. Then the average throughputs of the MSTFP protocol and TCP connections are respectively defined as follows:

$$T_M = \frac{\sum_{i=1}^{k_m} T_i^m}{k_m} \quad \text{and} \quad T_T = \frac{\sum_{i=1}^{k_t} T_i^t}{k_t}. \quad (16)$$

As such, the “friendliness” ratio can be defined as: $F = T_M / T_T$. (17)

[0074] To verify “rate smoothness” for the MSTFP protocol connections, let

$R_{m_i}^1, R_{m_i}^2, \dots, R_{m_i}^{S_{m_i}}$ represent the sending rates at different time instances 1, 2, ..., s_m of the i th MSTFP protocol connection and $R_{t_k}^1, R_{t_k}^2, \dots, R_{t_k}^{S_{t_k}}$ represent those of the k th TCP

connection, then the sending variation of the MSTFP protocol connections and the TCP connections are respectively defined as:

$$\Delta_{M_i} = \sum_{j=1}^{sm} |R_{m_i}^j - R_{m_i}^{j-1}| \text{ and } \Delta_{T_k} = \sum_{j=1}^{st} |R_{t_k}^j - R_{t_k}^{j-1}|. \quad (18)$$

[0075] The smoothness ratio is defined as

$$S = \Delta_{M_i} / \Delta_{T_k}. \quad (19)$$

[0076] By way of example, when $S \geq 1$ then the k th TCP connection is smoother than the i th MSTFP protocol connection.

[0077] Figure 5 shows the simulation results of the throughput as measured in Kb with respect to time measured in seconds for six (6) different connections. The data points for Figure 5 are listed in the Appendix. Two connections were used in an implementation of the MSTFP protocol (MSTPF1, MTPF2), two connections were the TCP protocol (TCP1, TCP2), and two connections (RAP1, and RAP2) were the Real time Adaptive Protocol (RAP) taught by Rejaie et al [SIGCOM]. The sending rates were measured in bps with respect to time. The six (6) different connections are illustrated in Figure 6. The data points for Figs. 5 and 6 are listed in the Appendix, below. From Figs. 5 and 6 it can be seen that the MSTFP protocol disclosed herein is more ‘friendly’ to TCP than RAP and also smoother than TCP and RAP.

[0078] The simulation used the MPEG-4 test sequence “*coast guard*” to illustrate the performance of the resource allocation scheme. The coast guard test sequence includes four objects – *river (obj0)*, *boat (obj1)*, *yacht (obj2)*, and *riverbank (obj3)*. The four objects were coded in CIF at a temporal resolution of 10 fps. In the simulation, the quality-impact parameter for these objects was selected as 0.7, 1.2, 1.2, and 0.7, respectively. The bottleneck link bandwidth (kb) is set to 100, 140, 100, 80, 100, 140,

100, 80, 100, and 140 to respectively correspond to time (ms) 0, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4500, and 5000.

[0079] The test sequence plotted in Figure 7, the data points for which are listed in the Appendix, shows the PSNR per frame of different objects. The thick line illustrates the resource allocation control scheme and the thin line is an illustration without the network resource allocation control scheme. Fig. 7 illustrates coding at CIF 7.5 Frames Per Second (FPS) using the MPEG-4 test sequence “Coast Guard” but, for the benefit of clarity, illustrates only the simulation results of *obj1* and *obj3*.

[0080] Figure 8 shows plots of the PSNR per frame coded at CIF 7.5 Frames Per Second (FPS) and using the MPEG-4 test sequence “Coast Guard”. The data points for Fig. 8 are listed in the Appendix, below. An implementation of the resource allocation control scheme is seen in a solid line and the dashed line illustrates a rate control scheme without this implementation. Note that there are several sharp drops in the resource allocation control scheme due to frame skipping. Furthermore, it can be seen that during the period of no frame skipping, the resource allocation control scheme has a higher PSNR. It can be seen in Figure 8 that, even taking the skipped frames into consideration, the implementation of the resource allocation control scheme achieves a better video quality. The data points for Figure 8 are listed in the Appendix, attached hereto.

[0081] Table 1 shows a comparison of the results with and without the MSTFP protocol resource allocation scheme disclosed herein. The multiple video streams for the MPEG-4 test sequence “*coast guard*” were used to produce the results seen in Table 1 for each of the four objects – *river (obj0)*, *boat (obj1)*, *yacht (obj2)*, and *riverbank (obj3)*.

[0082] **Table 1**

Resource control scheme	Obj0		Obj1		Obj2		Obj3		Average PSNR
MSTFP	W	PSNR	W	PSNR	W	PSNR	W	PSNR	24.16
Resource allocation	0.7	24.64	1.2	19.52	1.2	19.91	0.7	23.99	
Without control	PSNR		PSNR		PSNR		PSNR		20.99
	24.23		15.45		18.12		22.30		

[0083] 4. NETWORK ADAPTIVE RATE CONTROL SCHEME

An implementation of a network adaptive rate control scheme will now be presented in conjunction with the example illustrated in Figure 9. Also, the results of a simulation of an example of a network adaptive rate control scheme are set forth.

[0084] A. Discussion

The Network Adaptive Rate Control (NARC) scheme disclosed herein achieves minimal distortion in a global scene. Specifically, an overview is first given of an implementation of the NARC scheme, followed by a focus upon a quantizer adjustment of the NARC scheme.

[0085] Figure 9 shows a block diagram of the NARC scheme for multiple video objects (MVOs). The block diagram depicts a flow chart for the calculation of an estimate of the MSTFP available bandwidth that is made at module 910 of Figure 9.

[0086] Unlike the prior art rate control schemes taught by Chiang et al. and Vetro et al., implementations of the MVOs NARC scheme includes features such as dynamically estimating the available bandwidth at module 910, the target rate and global buffer

updating at module 912, quantization adjustment at module 934, and frame skipping control at module 960. Functionalities of these modules are discussed below.

[0087] The MVOs NARC scheme illustrated in Fig. 9 includes five stages, each of which will now be described. In the first stage, illustrated generally at module 910, the current available network bandwidth (R_T) is estimated in accordance with the MSTFP protocol. Together with the number of bits spent in the previous time instant (B_{prev}), the size ($R_{old}/2$) and the occupancy (W_{prev}) of the encoder output buffer, the target rate and global buffer are then updated for each Video Object Plane (VOP_i) frame at module 912. The output total target bits from the joint buffer control are allocated among each Video Object Plane (VOP_i) at module 914 to yield the target bits for each individual object. As the buffer size, $R_{old}/2$, is changed to $R_T/2$, the occupancy of the buffer, W_{cur} , is changed as follows:

$$W_{cur} = \max((W_{prev} + B_{prev}) \times R_T / R_{old} - R_T / F), 0, \quad (20)$$

where “F” is the video frame rate (e.g., 20 frame/second, 30 frame/second, etc.)

[0088] In the second stage, which is the pre-encoding part seen generally at reference numeral 930, the quantizer of each VOP_i is adjusted. Chiang et al. teach that the relationship between the quantization parameter (QP) and the texture value can be described as

$$r = \frac{p_1 \times MAD}{QP} + \frac{p_2 \times MAD}{QP^2}, \quad (21)$$

[0089] where p_1 and p_2 are control parameters. In the pre-coding part, initialization of the control parameters p_1 and p_2 takes place at initialization module 932. In order to maintain a constant quality of the total video sequence, the QP of each VOP_i is limited to

the range from 1 to 31, depending upon the bandwidth. The QP is allowed to change within a pre-defined range compared with the previous QP. A QP of 1 reflects the best quality and 31 represents the worst quality.

[0090] The Quantization Adjustment module 934 performs functions that contribute towards a global reduction in distortion. To achieve the global minimal distortion among the MVOs, the implementation of the MVOs NARC scheme employs the quadratic function to predict the number of bits, r_i , which will be used to encode the i th VOP:

$$r_i = \frac{(p_1)_i \times MAD_i}{QP_i} + \frac{(p_2)_i \times MAD_i}{QP_i^2}. \quad (22)$$

[0091] Still another quadratic function proposed by Chiang et al. is used to estimate the distortion, and is given by:

$$d_i = (q_1)_i \times QP_i + (q_2)_i \times QP_i^2, \quad (23)$$

where q_1 and q_2 are control parameters.

[0092] Verscheure et al. teach that packet-loss has an impact on video quality. A rate control implementation disclosed herein takes into consideration the packet-loss that occurs during video transmission. Video quality deteriorates quickly as the packet-loss rate and the sending rate go up.

[0093] To take the packet-loss into account, Eq. (23) is modified to:

$$d_i = (q_1)_i \times QP_i + (q_2)_i \times QP_i^2 + (q_3)_i \times r_i \times (P_L)_i, \quad (24)$$

where q_3 is the additional control parameter.

[0094] The objective of a rate control is to minimize the overall distortion to obtain the optimized quality for the whole video frame. Such a minimization can be expressed as:

$$\text{minimize } D = \sum_i d_i, \text{ subject to } R = \sum_i r_i \leq R_T, \quad (25)$$

where R_T is the total bit budget for the current time instant obtained from the buffer control algorithm and the MSTFP protocol. The third stage, seen generally in Fig. 9 at reference numeral 940, performs an encoding process based upon the quantization parameter (QP) estimate from the second stage, above. The “Intra coding” at module 944 receives the QP from pre-encoding Quantization Adjustment module 934 and is directed to the I-picture frame. The Quantization Adjustment module 934 is discussed further below. The “Inter coding” at module 942 receives initialized values from a pre-coding initialization module 932 and is directed to the P-picture frame or the predicted picture. Both “Intra coding” module 944 and the “Inter coding” at module 942 pass I-picture frames and P-picture frames, respectively, to the post encoding fourth stage seen generally at reference numeral 952 in Fig 9.

[0096] In the fourth stage, the rate model is updated at module 954. The objective of a rate control is to minimize the overall distortion. The update to the distortion model is performed at module 956. The rate and distortion models are updated based on the encoding results of the current frame as well as those from a specified number of the previous frames. Both the rate and the distortion aspects of the rate and distortion models are functions of the quantization parameter (QP) from Eq. 21 and the probability that the trace is in the loss state (P_L) from Eq. 4. The update to the rate and distortion models can be done using the least squares procedure described by Eckert et al.

[0097] The fifth stage is a frame skipping module 960 in Fig. 9. In the fifth stage, frame skipping occurs after all the video objects (VOs) in the scene have been encoded. Frame skipping occurs at this point in the implementation because, when the global scene is

considered, all the VOs composed within the same scene should be encoded at the same frame rate. For the purpose of frame skipping, the P-picture frame has a lower priority and the I-picture frame has a high priority. As such, the P-picture frame will preferably be the frame that is chosen to be skipped.

[0098] B. Simulation Results

To illustrate the bandwidth adaptation ability of an implementation of the NARC scheme, a simulation was conducted of the MSTFP protocol using the Network Simulator (NS) software version 2. The results of the simulation are illustrated in Figure 10 and the data points therefore are listed in the Appendix. Fig. 10, discussed below, demonstrates that given a variable available bandwidth, the MSTFP protocol can dynamically adjust the sending rate to match up with the available network bandwidth.

[0099] Figure 10 depicts the available network bandwidth and its estimation in plot of bandwidth measured in kbps against time measured in ns. The available network bandwidth is illustrated in Fig. 10 as a dashed line and the MSTFP estimated network bandwidth is illustrated by a solid line.

[0100] In the simulation, a standard MPEG-4 codec was used with unrestricted motion vector modes. The simulation tested two rate control schemes: 1) the bandwidth adaptive MVOs rate control scheme disclosed herein, and 2) a conventional MPEG-4 standard MVOs rate control (Q2) without any feedback from network bandwidth and packet-loss rate. A change was made in the frame skipping part of the conventional MPEG-4 Q2 scheme so as to maintain all the objects of the same scene in a consistent frame rate. In both cases, the first frame was intracoded, and the remaining frames were intercoded. The tests used the MPEG-4 “coast guard” video test sequence and the MPEG-4 “news” test video sequences, respectively. Both of them are composed of four video objects.

[0101] PSNR was used in the simulation as a metric to measure video quality. For an eight-bit image with intensity values between 0 and 255, the PSNR is given by:

$$PSNR = 20 \log_{10} \frac{255}{RMSE}, \quad (26)$$

where RMSE stands for root mean squared error. Given an original $N \times M$ image f and a compressed or degraded image f' , the RMSE can be calculated as follows:

$$RMSE = \sqrt{\frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [f(x,y) - f'(x,y)]^2}. \quad (27)$$

[0102] A fair comparison of PSNR between the two rate control methods with the skipping of different frames is not trivial. A rate control technique that skips more frames would typically spend more bits per coded frame and could easily have a very high average PSNR per coded frame. Table 2, below, presents the rate control impact on video quality. The average PSNRs of the encoded frames are shown in the third column of Table 2. It can be seen from Table 2 that there is more frames skipping in the conventional MPEG-4 Q2 scheme than in an implementation of the NARC scheme.

[0103] **Table 2**

Video Sequence	Rate control scheme	Encoded average PSNR	Skipped frame	Total average PSNR
Coast guard (~100 kbps)	NARC scheme	26.02	20	23.92
	MPEG-4 Q2 scheme	27.33	26	21.89

News (~100 kbps)	NARC scheme	30.97	8	30.77
	MPEG-4 Q2 scheme	33.08	25	29.36

[0104] Table 2 also demonstrates that the encoded average PSNR in the conventional MPEG-4 Q2 scheme is higher than in the NARC scheme. However, using the encoded average PSNR value as a measure of compressed video quality is not accurate. This is because the distortion of the non-encoded video is not taken into account. In the rate-control testing in the conventional MPEG-4 Q2 scheme, it was decided that when a frame was skipped, the previous encoded frame should be used in the PSNR calculation because the decoder displays the previous encoded frame instead of the skipped one. The average PSNR is calculated using this approach herein and the results are shown in Table 2 as the “TOTAL average PSNR” in column 5.

[0105] Figs. 11 and 12 each show a plot of PSNR per frame at 10 frames per second, where Fig. 11 uses the MPEG-4 “coast guard” test sequence and Fig. 12 uses the MPEG-4 “news” test sequence. The respective data points for each of Figures 11 and 12 are listed in the Appendix. In Figures 11 and 12, the solid line illustrates an implementation of the NARC scheme and the dashed line illustrates the convention MPEG-4 Q2 scheme. The PSNR values of the tested video drops sharply due to frame skipping. Figs. 11 and 12 illustrate that the NARC scheme significantly reduces the number of frames that are skipped. During the period when no frames are skipped in either method, the NARC scheme has a higher PSNR, and therefore video quality, than that of the conventional

MPEG-4 Q2 scheme. This is because the packet-loss rate in the NARC scheme is less than that in the conventional MPEG-4 Q2 scheme.

[0106] The above-described implementations of systems and methods are presented herein in the general context of server-client architecture via the Internet. The client and server may be implemented by general-purpose computers, or by other particularly tailored computing devices, such as minicomputers, mainframe computers, workstations, Internet appliances, set top boxes, wireless communication devices, game consoles, portable computing devices, and the like.

[0107] An exemplary computer system has one or more processors, memory, and one or more input/output (I/O) devices. The processor communicates with the memory and I/O devices via one or more bus/interface mechanisms. Programs, data, files, and the like may be stored in memory and executed on the processor.

[0108] The memory typically includes both volatile memory (e.g., RAM) and non-volatile memory (e.g., ROM, Flash, hard disk, CD-ROM, DVD, etc.). Examples of possible I/O devices include a mouse, a touch pad, a joystick, a keyboard, a keypad, a display, a peripheral storage device, and a microphone. The computer system may also implement a computer operating system that is stored in memory and executed on the processor.

[0109] The present invention may be embodied in other specific forms without departing from its spirit or essential characteristics. The described embodiments are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is, therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.

[0110] APPENDIX

Fig. 5 – Six (6) plots of sixty (60) data points, from 1 to 60:

<u>MSTFP1</u>										
25	79	88	124	125	155	153	183	175	209	197
228	218	248	239	271	255	289	277	312	302	336
324	357	346	381	366	403	389	424	413	447	431
467	451	489	471	508	492	529	511	549	529	568
552	588	572	607	587	625	606	644	626	664	646
684	667	703	688	725						

<u>MSTFP2</u>										
14	29	42	53	67	78	92	101	115	123	135
143	156	161	176	184	192	201	209	218	231	239
248	258	268	279	287	298	308	315	329	338	348
355	366	376	387	397	409	420	426	440	446	461
470	484	493	507	513	528	533	548	555	570	577
594	600	615	621	640						

<u>TCP1</u>										
68	36	85	50	87	58	123	103	162	118	183
141	198	149	239	169	261	205	297	253	330	261
365	286	403	304	415	339	443	361	481	374	492
423	531	427	558	471	584	495	600	533	628	542
643	593	656	596	662	614	696	640	734	677	756
680	756	680	756	681						

<u>TCP2</u>										
31	37	41	57	68	93	103	99	119	105	127
117	152	138	175	145	183	154	204	164	216	171
248	183	252	213	296	219	302	220	304	220	310
227	323	257	340	292	375	308	393	345	431	404
477	411	506	463	556	474	568	508	590	539	610
586	643	634	690	639						

<u>RAP1</u>										
1	1	63	63	104	113	136	169	176	234	240
273	273	317	308	383	372	422	401	473	451	526
501	571	538	607	586	665	633	730	684	794	771
849	827	912	858	939	906	982	963	1031	999	1064
1048	1116	1105	1155	1159	1220	1216	1260	1253	1292	1282
1341	1337	1384	1394	1455						

<u>RAP2</u>										
1	1	5	24	60	62	91	89	133	141	184
257	257	286	305	326	364	389	427	427	473	501
505	566	570	625	614	673	676	734	727	789	773
828	813	872	884	914	919	975	979	999	1026	1046
1056	1081	1101	1133	1143	1186	1184	1235	1232	1281	1306
1325	1357	1381	1391	1435						

[0111] Fig. 6 - Six (6) plots of sixty (60) data points, from 1 to 60:

MSTFP1

68022	42745	14326	16720	17420	22844	24124	22042	25547	21249	23897
20511	19820	20891	18771	17810	20177	19883	18367	23842	26506	22961
20556	21740	19676	20867	24021	19402	21519	18487	21310	23562	23279
20327	20748	26960	20312	27401	21843	25885	24942	20751	28371	23268
23941	27077	22757	25453	22119	22617	24809	26673	21646	19403	25080
22281	20460	21282	19322	22834						

MSTFP2

65151	38841	27793	26881	34153	22795	23252	20361	23289	22110	28267
24500	22071	19041	21058	22649	19497	20840	21043	18842	23085	20188
23489	21368	17728	17088	20880	19863	21691	18292	17592	19425	23435
17840	19008	16702	25920	16272	16644	18879	18673	15109	20391	16378
16803	15364	18652	16139	21319	18891	21805	19392	15544	14734	17505
18399	15959	18110	17044	19357						

TCP1

7812	16815	26785	37640	14423	5750	62500	26785	9740	31250	18700
8690	9810	12560	23810	34200	21200	12390	7690	8780	18729	36840
20120	8920	16802	34280	36901	13920	7690	16590	32760	17650	4329
8903	24390	43290	30912	10098	7654	12458	30190	42190	28760	12310
7630	12865	43210	21650	9860	18702	34908	24190	10087	6530	15490
32980	28750	11980	8750	15439						

TCP2

16815	16815	20833	43750	8928	8522	12500	7500	56000	12416	7900
35900	23001	8790	10029	16709	27650	13280	6540	6508	14890	31290
18769	9870	18760	30920	26340	11890	5670	15320	31240	15609	4356
10092	15430	32900	24319	12390	6542	14230	28971	36570	25430	13290
5430	16430	39870	23190	10020	16540	28760	19874	7860	9970	18760
36740	24130	13540	6530	18453						

RAP1

5890	45690	52806	43952	34898	46980	61128	57247	41099	34270	36179
60945	34002	8900	9200	24540	43250	21340	9657	12680	18540	34200
38790	36570	28769	21560	13450	7865	16540	19870	27870	34526	42536
39870	31260	24670	16570	9345	6349	23950	34281	42536	54390	43120
32120	16230	8740	18670	32454	42356	31290	12987	8766	18760	36542
18726	8765	16540	35480	12980						

RAP2

8730	10040	17609	36844	51746	60045	30805	81163	41657	77964	77040
52345	20450	10030	12420	32150	37980	54370	8790	18650	22980	29870
34526	33456	26574	20980	16598	9675	18657	20980	24678	38970	48790
41230	34350	21809	13460	8234	10049	28730	35640	48670	49830	39820
28120	16532	6450	16540	29810	39120	28718	14320	7658	20932	43290
15234	6540	19870	38708	21543						

[0112] Figure 7 – Four (4) plots of thirty-eight (38) data points, from zero to 37:

<u>Obj1 (No)</u>										
30.37	25.3	19.92	19.9	19.62	19.4	20.99	18.82	17.36	17.09	16.66
16.93	15.31	14.11	13.97	13.69	13.73	14.45	13.44	12.57	12.58	12.69
12.71	12.77	12.8	12.89	13.12	13.13	13.09	13.05	12.95	12.88	12.9
13	13.11	13.16	13.26	13.37						

<u>Obj3 (N0)</u>										
31.46	30.98	30.2	29.81	29.6	29.28	29	28.7	28.53	28.04	27.51
27.23	25.43	23.75	23.27	22.87	22.46	22.82	17.6	16.41	16.44	16.58
17.16	17.37	17.37	17.45	17.65	17.67	17.55	17.74	17.92	18.18	18.3
18.49	18.9	19.14	19.26	19.44						

<u>Obj1</u>										
30.37	8.51	7.03	6.7	25.11	24.35	24.23	23.78	23.82	23.72	24.21
24.07	24.21	8.68	24.23	8.56	6.75	24.88	4.09	24.33	9.98	24.09
24.03	23.99	24.02	23.84	23.83	24.22	12.42	24.23	11.26	9.13	24.08
14.88	23.94	23.94	24.14	24.26						

<u>Obj3</u>										
31.46	20.79	19.94	19.66	26.47	27.35	27.34	27.49	27.54	27.61	27.73
27.83	27.15	18.87	26.8	18.61	17.57	27.79	14.97	25.51	18.73	26.06
24.98	25.62	25.9	25.97	26.12	25.23	16.96	25.85	16.74	16.19	25.76
18.34	25.93	26.38	26.4	26.2						

[0113] Figure 8 – Two (2) plots of thirty-eight (38) data points, from zero to 37:

<u>With Resource Allocation</u>										
31.38	22.81	21.44	20.69	26.68	26.88	26.82	26.7	26.68	26.61	26.67
26.66	26.43	19.71	26.3	19.18	16.99	27.67	15.57	25.99	18.18	26.17
25.51	25.82	26.02	26	26.2	25.9	18.1	26.2	17.28	16.53	26.21
20.29	26.24	26.58	26.64	26.57						

<u>Without Resource Allocation</u>										
31.36	28.04	27.36	26.93	26.79	26.53	26.61	25.88	24.87	24.46	23.97
23.81	22.46	21.37	20.97	20.68	20.54	21.5	17.26	16.28	16.32	16.58
16.97	17.12	17.22	17.32	17.64	17.8	17.51	17.61	17.66	17.88	18.08
18.3	18.64	18.88	19.04	19.34						

[0114] **Figure 10 - Two (2) plots of fifty (50) data points, from 0.0 to 4.9, in increments of 0.1:**

Available Bandwidth

100	100	100	100	100	140	140	140	140	140	100
100	100	100	100	80	80	80	80	80	100	100
100	100	100	140	140	140	140	140	100	100	100
100	100	80	80	80	80	80	100	100	100	100
100	140	140	140	140	140					

MSTFP Available Bandwidth

98	103	99	102	98	136	141	144	138	142	103
99	105	97	101	83	80	77	81	79	98	102
99	103	97	136	142	144	139	136	102	99	102
98	95	81	76	79	82	79	98	103	100	95
98	137	140	143	141	138					

[0115] **Figure 11 – Two (2) plots of fifty (50) data points from 1 to 50, incrementing by 1:**

NARC Scheme

31.36	23.47	21.96	21.2	20.67	20.16	28.16	21.94	27.16	26.93	26.81
26.76	26.72	26.73	26.84	26.71	20.45	26.35	20.18	18.41	26.14	18.95
17.08	28.6	15.98	26.32	19.73	18.81	26.27	26	26.29	26.41	26.44
26.47	26.54	26.78	26.57	18.86	26.26	18.75	16.68	16.32	26.3	20.83
19.23	26.29	26.98	27.11	27.15	27.17					

MPEG-4 Q2 Scheme

31.36	28.38	28.1	28.07	27.91	27.9	27.79	27.74	27.81	27.65	27.39
27.16	27	27.04	27.07	25.62	25.62	20.54	24.41	20.06	18.45	17.23
25.32	17.12	15.14	14.1	26.83	20.86	17.87	17.39	16.8	16.41	15.96
25.21	19.56	17.48	16.24	15.84	15.49	22.82	17.85	17.05	16.67	16.53
23.91	19.12	17.36	16.88	16.89	20.37					

[0116] Figure 12 – Two (2) plots of seventy-four (74) data points, from 1 to 74, incrementing by 1:

NARC Scheme

34.76	28.22	26.89	25.72	25.38	25.11	24.65	24.34	24.06	33.68	33.41
32.95	32.64	32.35	32.26	32.18	32.04	31.93	32.03	32.04	32	31.94
31.97	31.81	31.94	31.85	31.93	31.86	31.76	31.67	31.65	31.51	31.6
31.65	31.64	31.51	31.48	31.49	31.35	31.25	31.14	31.03	31.06	30.86
30.84	30.85	30.61	30.69	30.76	30.78	30.93	30.84	30.84	30.75	30.79
30.56	30.61	30.66	30.63	30.75	30.73	30.75	30.83	30.78	30.75	30.55
30.63	30.59	30.77	30.78	30.65	30.74	30.93	30.83	30.9		

MPEG-4 Q2 Scheme

34.76	28.22	26.89	25.72	33.93	33.58	33.06	30.85	32.62	32.6	28.08
32.29	32.32	27.83	32.08	28.03	31.83	31.84	31.87	31.9	31.83	27.72
31.83	31.61	31.6	31.59	31.52	31.41	31.44	31.34	31.31	27.39	31.25
31.4	31.5	31.36	27.19	31.26	31.32	28.8	31.08	31.07	27.56	30.88
27.53	30.93	22.66	30.58	25.03	30.66	23.99	30.76	26.52	27.71	24.92
27.39	23.9	27.13	27.5	28	28.16	25.83	28.45	28.41	28.57	24.52
28.74	24.63	29.18	29.35	23.48	29.41	24.99	30.01	27.29		